REGRESSION TECHNIQUES FOR THE CALCULATION OF BALLISTIC WINDS UTILIZING SATELLITE INFRARED SPECTROPHOTOMETER DATA

John Richard Wright

United States
Naval Postgraduate School

THESIS

REGRESSION TECHNIQUES
FOR THE CALCULATION OF BALLISTIC WINDS
UTILIZING SATELLITE INFRARED SPECTROPHOTOMETER DATA

by

John Richard Wright

Thesis Advisor: Russell L. Elsberry

September 1971

Regression Techniques
for the Calculation of Ballistic Winds
Utilizing Satellite Infrared Spectrophotometer Data

by

John Richard Wright
Lieutenant, United States Navy
B.S., University of Missouri, 1964

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY

from the

NAVAL POSTGRADUATE SCHOOL
September 1971

ABSTRACT

Three regression techniques for the calculation of ballistic winds using satellite infrared spectrophotometer (SIRS) data are evaluated. Experiments testing the derived methods are carried out using two SIRS-B data set. In all three cases, the data sets are divided into a large dependent sample used to derive the regression equations and smaller independent sample for testing the equations' validity. One method, that of geostrophic calculations using regression derived "D" values of ballistic geopotential height is found to be superior to the other methods.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

ACKNOWLEDGEMENTS

The author offers his grateful thanks to Dr. Russell L. Elsberry for his constant guidance, endless patience, and occasional necessary prodding during the course of this research. Also, a note of appreciation to the staff of the W. R. Church Memorial Computer Center with special thanks to William D. Ehrman for his invaluable assistance.

I. INTRODUCTION

The resultant wind which deflects a reentry vehicle as it travels along its flight path is known as the ballistic wind. This is a weighted integral of the winds encountered at the various levels of the atmosphere through which the projectile travels. The general form of the ballistic wind calculation is:

$$\mathbb{W}_B = \frac{1}{\ln p_{oo}} \int_0^{p_{oo}} \mathbb{V}(\ln p)\, W_v(\ln p)\, d(\ln p) \qquad (1)$$

where $\mathbb{W}_B$ is the ballistic wind, p is pressure, p_{oo} is surface pressure, $\mathbb{V}$ is the actual wind at level p, and W_v is the ballistic wind weighting factor for a layer. W_v is proportional to the effect of the winds in a particular layer and the time the vehicle remains in that layer. A similarly computed quantity, ballistic density, is representative of the total retarding force of the atmosphere on a reentry vehicle.

Present methods of evaluation of ballistic winds and density involve the use of rawinsonde data and rocketsonde data as available for the levels in the integration of Eq. (1). Smith, et. al. [6] have recently reported success in obtaining atmospheric temperature-pressure profiles using satellite infrared spectrophotometer (SIRS) data. This success implies that ballistic winds and densities could be calculated in any region without the necessity of a rawinsonde release.

Ruggles [5] has noted that the envelope of weighting functions used with the individual SIRS channel radiances in obtaining atmospheric soundings and the ballistic wind weighting factors are analogous. He further suggested that the ballistic winds and densities might be calculated directly from SIRS data without first calculating the temperature-pressure profile. Ruggles' development is outlined below.

It is the object of this research to investigate the feasibility of Ruggles' suggestion with respect to ballistic winds.

II. DEVELOPMENT OF EQUATIONS

A. FEASIBILITY OF THE USE OF SIRS DATA

If the geostrophic approximation is assumed, Eq. (1) can then be expressed as:

$$\mathbb{V}_B = \frac{1}{\ln p_{oo}} \int_0^{p_{oo}} -\frac{g}{f} \, \mathbb{k} \times \nabla_p Z(x, y, \ln p) W_v(\ln p) \, d(\ln p) \tag{2}$$

or, rearranging terms:

$$\mathbb{V}_B = \frac{-g}{f \ln p_{oo}} \, \mathbb{k} \times \nabla_p \int_0^{p_{oo}} Z(x, y, \ln p) W_v(\ln p) \, d(\ln p) \tag{3}$$

where g is the acceleration due to gravity, f is the Coriolis parameter, and ∇_p is the horizontal gradient of geopotential height on a constant pressure surface. Now, expressing the weighting function W_v as a linear combination of known arbitrary functions of ln p, F_i (ln p):

$$W_v (\ln p) = \sum_{i=1}^{n} a_i F_i (\ln p) \tag{4}$$

where a_i are constants. Similarly, Z(x, y, ln p) can be expressed as:

$$Z(x, y, \ln p) = \sum_{i=1}^{n} r_i H_i (x, y, \ln p) \tag{5}$$

where r_i are the constants and H_i are a set of arbitrary functions of ln p. Substitution of Eq. (4) and Eq. (5) into Eq. (3) produces:

$$\mathbb{V}_B = \frac{-g}{f \ln p_{oo}} \, \mathbb{k} \times \nabla_p \sum_{i=1}^{n} a_i r_i \int_0^{p_{oo}} H_i(x, y, \ln p) \, F_i(\ln p) \, d(\ln p) \tag{6}$$

Now consider the radiative transfer equation for the radiance observed by the i^{th} channel of the SIRS, sensing at a wave number, ν .

$$N_i = N_{o_i}[A, \nu, T(\ln p), \tau(\nu, \ln p)] + \int_{p_{oo}}^{0} B_i[\nu, T(\ln p)] \, \frac{d\tau_i(\nu, \ln p)}{d(\ln p)} \, d(\ln p) \tag{7}$$

Here, N_i is the radiance sensed, and N_{0_i} is the radiance emitted by a contaminant such as the top surface of clouds. This N_{0_i} is a function of A, the amount of cloud cover, ν , the wave number of the i^{th} channel, and T(ln p) the temperature structure above the clouds. The Planck radiance at wave number, ν , and temperature, T, is represented by $B_i[\nu, T(\ln p)]$, and $\tau(\nu, \ln p)$ is the fractional transmittance of the atmosphere at wave number ν from pressure level p to the SIRS.

Smith et. al. [6] have shown that cloud contaminant term can be evaluated by an iterative procedure. With that term evaluated, Eq. (7) becomes:

$$N^*_i = \int_{p_{oo}}^{0} B_i[\nu, T(\ln p)] \, \frac{d\tau_i(\nu, \ln p)}{d(\ln p)} \, d(\ln p) \tag{8}$$

where N^*_i is the radiance corrected for cloud contamination effects, also called the clear column radiance. Since the only constraints on F_i are that they are known functions of ln p, we can specify, (after Ruggles [5]):

$$F_i = \frac{d\tau_i(\nu, \ln p)}{d \ln p}$$

Similarly, since $B_i[\nu, T(\ln p)]$ is an increasing function of T and also of ln p, it satisfied the constraints put on H_i in Eq. (5) and we can set:

$$H_i = B_i[\nu, T(\ln p)]$$

Substituting these expressions into Eq. (8) produces:

$$N_i^* = \int_{p_{oo}}^{0} H_i F_i \, d(\ln p) \tag{9}$$

Using this in equation (6) results in a direct expression for ballistic winds in terms of the SIRS radiances.

$$\mathbb{V}_B = -\frac{g}{f \ln p_{oo}} \mathbb{k} \times \nabla_p \sum_{i=1}^{n} a_i r_i N_i \tag{10}$$

Now n is the number of SIRS channels used. In practice, the constant terms are grouped together and are statistically determined through the use of the BIMED 02R stepwise multiple regression routine, (Dixon, 1) available at the W. R. Church Computer Center at the Naval Postgraduate School. This produces a final equation of the form:

$$\mathbb{V}_B = -\frac{1}{f} \mathbb{k} \times \nabla_p \left[\sum_{i=1}^{n} b_i N_i + C_1 \right] \tag{11}$$

where b_i are the regression coefficients and C_1 is the regression equation constant. The constant terms, g and ln p_{oo} are absorbed by the coefficients, b_i.

B. "D" VALUE APPROACH

The bracketed term in Eq. (11) represents the regression equation. If the ballistic wind, V_B, is geostrophic, this term can be looked upon as a ballistic geopotential height. The regression equation can be derived by correlation of SIRS radiances with a dependent sample of ballistic geopotential heights. These are obtained by computing a weighted average of observed geopotential heights in the vertical. The ballistic wind weighting factors are used in the averaging. For the dependent sample, deviations from standard, or "D" values of height were used for all available mandatory levels from rawinsonde soundings. For levels above the maximum height attained by the rawinsonde (usually 10 mb) temperature values were extrapolated using the standard lapse rate for 45N in January. The hydrostatic approximation was then used to calculate "D" values. Linear interpolation was used in the rare instances when an interior mandatory level value was missing in a sounding. The following equation illustrates the method used for obtaining the dependent sample of "D" values.

$$D_B = \sum_{p=1000}^{1\,mb} W_p D_p \qquad (12)$$

Here, D_B is the ballistic "D" value, W_p are the ballistic wind weighting factors for each mandatory level given by Finke [3], and D_p are the actual "D" values of height at each mandatory level.

This method of calculating V_B, then, consists of two steps. First a regression equation is derived by correlation of SIRS radiance

data with the dependent sample of "D" values calculated using Eq. (12). Then, Eq. (11) is used with the derived regression equation to calculate the ballistic wind.

C. GRADIENT OF RADIANCES APPROACH

Rearrangement of the order of Eq. (11) yields

$$\mathbb{V}_B = \frac{-1}{f}\left[\sum_{i=1}^{n} c_i \, \mathbb{K} \times \nabla_H N_i + C_2\right] \tag{13}$$

The regression coefficients are represented by c_i, and C_2 is the regression equation constant. Here, c_i and C_2 are different than b_i and C_i in Eq. (11). Though equivalent, this is procedurally different in evaluation from Eq. (11). In this case the entire expression is a regression equation. The equation is derived by correlation of values of $\frac{1}{f}\frac{\partial N_i}{\partial x}$ with v_B, the north-south component of $\mathbb{V}_B$. Similarly, u_B, the east-west component is correlated with $\frac{1}{f}\frac{\partial N_i}{\partial y}$. The dependent sample of u_B and v_B is obtained as before from rawinsonde soundings using Eq. (1). Again, missing interior mandatory level values were supplied by linear interpolation. Missing upper level values of u were extrapolated by paralleling a mean wind profile. Missing v values were extrapolated by simply decreasing the already small v component by five per cent with each level increment. This was justifiable since v components above 10 mb were small in all cases.

D. DIRECT APPROACH

A purely statistical method suggested by Dr. F. L. Martin was to perform direct correlations between the SIRS radiances and the ballistic

wind components. This produces a very simple regression equation for the ballistic winds.

$$\mathbb{W}_B = \sum_{i=1}^{n} h_i N_i + K \tag{14}$$

Here, h_i are the regression coefficients and K is the regression equation constant.

III. SIRS-A EXPERIMENT

A. DATA AND DATA HANDLING

The data used in the SIRS-A experiment were obtained from Project FAMOS and consisted of radiances, corrected for cloud contamination, for the eight SIRS channels for two four-day periods: December 27-30, 1969, and January 2-5, 1970. The area of data coverage was the European continent including the area of the U.S.S.R. west of the Urals, and the time period covered was from 0800 GMT to 1200 GMT for each day. The dependent ballistic "D" values and winds were computed from the 1200 GMT rawinsonde data from the same periods. The rawinsonde data did not include any data from France or the British Isles. In some cases this was a limitation due to the location of satellite subtracks lying in this data sparse zone. In addition, data was typically quite sparse in the Mediterranean region.

The 27 degree longitudinal spacing of Nimbus III SIRS data precludes the calculation of both u and v components of ballistic winds by the geostrophic approximation. Instead, V_{BN}, the wind component normal to the satellite subtrack was calculated using Eqs. (11) and (13) modified as indicated below

$$\mathbb{V}_{BN} = -\frac{g}{f}\frac{\partial}{\partial n}\left[\sum_{i=1}^{n} b_i N_i + C_1\right] \tag{11a}$$

$$\mathbb{V}_{BN} = -\frac{1}{f}\left[\sum_{i=1}^{n} d_i \frac{\partial N_i}{\partial n} + C_2\right] \tag{13a}$$

Here n is the generally northerly direction parallel to the satellite subtrack. The equation for the direct approach remained unchanged.

The "D" values and values of u_B and v_B calculated from all available rawinsonde data were plotted and each of the fields analyzed. Values were then interpolated to SIRS subtrack positions. Since the points with SIRS data were irregularly spaced, the data were processed using a cubic spline fitting interpolation routine which yielded radiance data at evenly spaced points at a 2 degree latitude interval along the satellite subtrack. The effects of this interpolation on the methods' accuracy are discussed later. A typical Nimbus III subtrack and data is shown in Fig. 1.

B. "D" VALUE APPROACH

The initial regression equations used January 2-4 data as a dependent data set. Initial experimentation was confined to the January data because the missing French and British data caused a bad gap in data coverage relative to the positions of SIRS data on December 29.

At first, two dependent samples of "D" values were correlated with radiance data, one using interpolated evenly spaced radiance data and another using data from the uninterpolated sub-points. Results using the two different data sets were similar and are summarized in Table 1. Since the difference in results was slight and the evenly spaced data were much easier to handle, experimentation proceeded with the interpolated data.

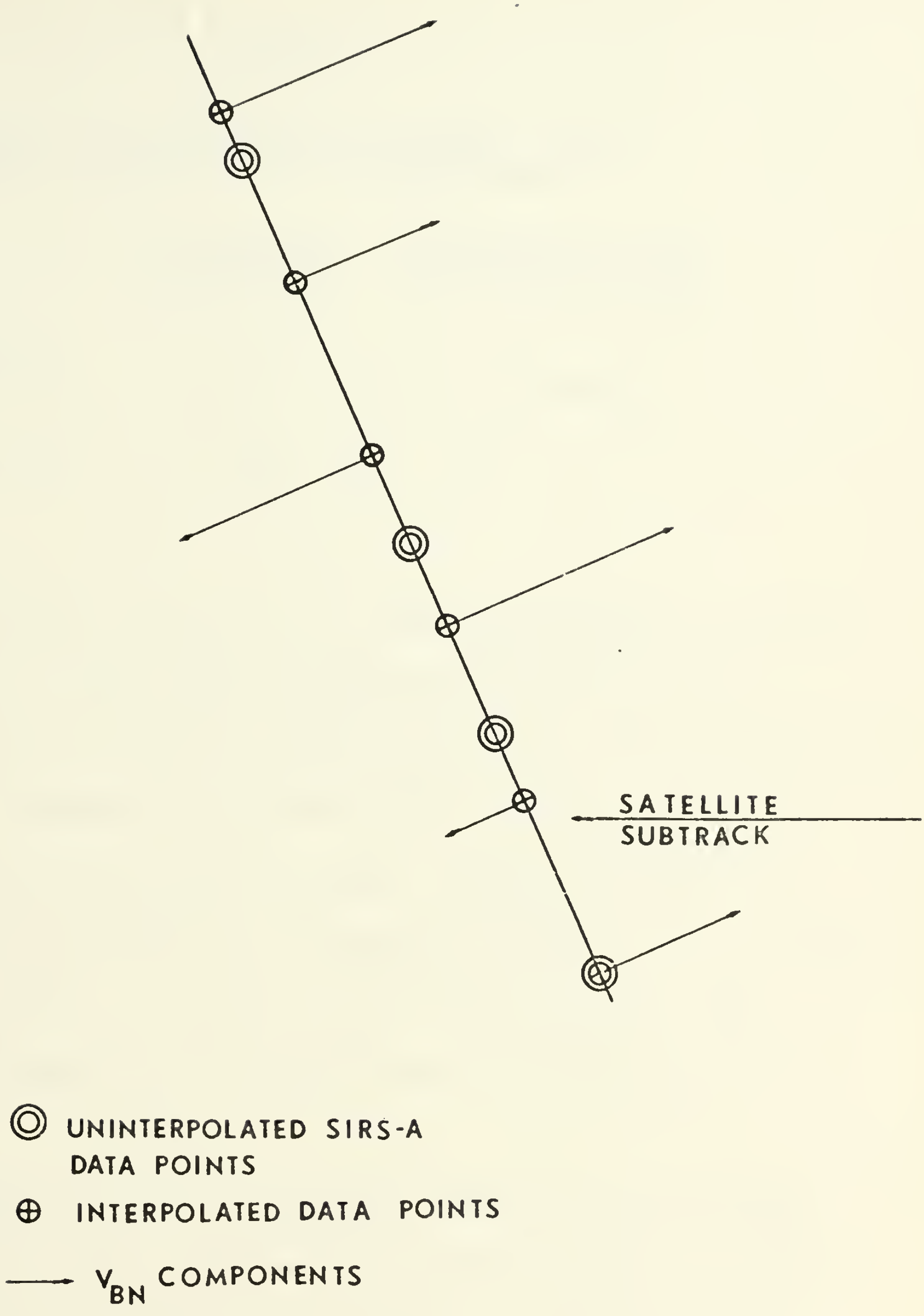

Fig. 1. Typical Nimbus III Satellite Subtrack Showing Uninterpolated and Interpolated SIRS-A Data.

TABLE I

Effect of Interpolation on Results of "D" Value Correlations with SIRS Radiances

	Correlation Coefficient	Standard Error of "D" Value Estimate
Interpolated Data	.975	51.6 meters
Uninterpolated Data	.979	49.7

Since the standard errors of "D" value determination were quite large compared to those found by Elsberry and Martin [2] for a dependent sample of simulated radiance data (22.3 meters), the data were carefully examined. It was noted that the great majority of points which had large residuals between rawinsonde derived "D" values and those obtained from the regression equation were located in data-poor regions. It was therefore assumed that the large residuals were mainly due to inaccuracies in the dependent "D" values.

In an attempt to reduce the standard errors of the estimated "D" values, those data points which had produced residuals greater than 50 meters were removed from the sample. For this filtered sample, a new correlation was calculated with somewhat improved results, but this reduced the size of the data set from n = 90 to n = 49. The correlation was increased to .987, and the standard error was reduced to 28.3 meters.

For the December data, the same two correlation runs were performed with the following results: for all data points (n = 80)

the multiple correlation coefficient was .806 and the standard error was 51.8 meters. For the filtered data ($n = 53$) the correlation was .946 and the standard error was 27.6 meters. The four regression equations for ballistic "D" value, D_B, are listed below:

December, unfiltered data:

$$D_B = 1.281N_1 - 5.243N_2 + 13.642N_3 + 7.884N_4 - 58.512N_5 + 63.779N_6 + 9.836N_7 - 1.088N_8 - 1822.8 \tag{15}$$

December, filtered data:

$$D_B = 1.212N_1 - 7.702N_2 + 14.870N_3 + 8.320N_4 - 58.870N_5 + 66.715N_6 + 12.54N_7 - 1.834N_8 - 1986.0 \tag{16}$$

January, unfiltered data:

$$D_B = -5.497N_1 - 58.457N_2 - 42.484N_3 + 32.944N_4 - 91.296N_5 + 40.878N_6 - 10.980N_7 + 14.980N_8 + 1792.3 \tag{17}$$

January, filtered data:

$$D_B = -5.622N_1 - 51.070N_2 + 28.321N_3 + 20.475N_4 - 50.746N_5 + 19.161N_6 + 1.928N_7 + 12.966N_8 + 1275.7 \tag{18}$$

A comparison of the magnitudes of the regression coefficients indicated the significance of the different SIRS channels varies markedly from the December to the January data sets. In December, channels 5 and 6 provide the major modification to the constant term while in the January case, channels 2, 3, 4, 5, and 6 all had contributions of roughly the same order of magnitude. The coefficient for channel 5, however, is almost twice as large as the coefficients

for the other channels. Results reported by Smith, et. al. [6] would indicate that the principal contribution in both cases came from the layer centered around 250-150 mb. In the January case, however, it would appear the principal contribution came from a slightly lower level and that the vertical wind shear was less.

That the regression equations derived for the filtered and unfiltered cases were nearly equivalent was demonstrated by comparing "D" values calculated from both periods. The comparisons are tabulated in Table II.

TABLE II

Comparison of "D" Values Obtained from Filtered and Unfiltered Regressions

	Correlation	Explained Variance	Std. error
Dec.	.993	98.6%	8.7 meters
Jan.	.995	99.0	20.5

To check this scheme for calculating usable ballistic winds, the "D" values calculated using all four of the derived regression equations, Eqs. (15) - (18) were used with Eq. (11b) to generate ballistic wind values. These were then compared with the rawinsonde derived values as summarized in Table III.

The <u>mean of the absolute values</u>, $\overline{|V_{BN}|}$, of the wind components is included in the table since both easterlies and westerlies may be included in the sample. In December, the winds were quite variable and $\overline{V}_{BN}$ and $\overline{|V_{BN}|}$ differ greatly. In January the winds were highly zonal and $\overline{V}_{BN}$ and $\overline{|V_{BN}|}$ are equal.

TABLE III

Comparison of Regression Equation-Derived and Rawinsonde-Derived Ballistic Winds Using "D" Value Approach (SIRS-A) for Independent Samples in Dec. and Jan.

Independent Sample	n	Mean V_{BN}	Std Dev of V_{BN}	Mean V_B	Corr Coeff	Explained Variance	Std Error	Max Error
Dec.								
$V1-V_B(I)$	23	4.7kts.	29.1	24.2kts.	.896	80.3%	13.3kts.	39.6kts.
$V2-V_B(I)$	23	"	"	"	.929	86.4	11.0	30.8
Jan								
$V1-V_B(I)$	11	40.0	8.2	40.0	.822	67.5	4.9	9.3
$V2-V_B(I)$	11	"	"	"	.802	64.3	5.2	10.9

Here, V1 represents ballistic winds calculated using the regression equation derived from all data points, and V2, those obtained from the equation derived from filtered data. The (I) indicates the independent samples, either December 30 or January 5. Although the filtered and unfiltered dependent equations were used, both were applied to the same, unfiltered data set for the test of the independent sample. An unfiltered data set was used as a check since it would be impossible to perform the filtering process on an independent data set if the procedure were to be used operationally.

It was found that the significance of the channels entering into the dependent regression equations, Eqs. (15 - 18), became very slight after more than five channels were used. The correlation coefficient was increased by less than 1.0 per cent with the addition of each succeeding channel. In spite of this, regression equations were derived using all SIRS channels since the addition of succeeding channels reduced the standard error.

It can be seen that the filtering of the data had little effect on the final outcome. In fact, the results were opposite in the two cases. For the independent December data, filtering slightly improved the results, while it slightly degraded them in January. In either case, the difference was insignificant, at most amounting to a 6.1 per cent difference in explained variance or a 2.3 knot difference in standard error. The standard errors of the wind component are not overly large considering the wind is a weighted average from the surface to 30 km. In any case, they are well within the accuracy criteria set down by Finke [3].

For both periods, the wind estimates obtained by using the geostrophic approximation with SIRS derived "D" values are better than those derived from "D" values obtained from the rawinsondes. Geostrophic values of V_{BN} were calculated using "D" values interpolated to the satellite subpoints and compared with the calculated values of V_{BN} at those points. A correlation of .725 was obtained for the December data and .766 was the correlation for January. An explanation of the superior performance of the SIRS techniques is that the "D" value fields obtained from SIRS data are free of the inaccuracies inherent in a field obtained by many different sensors. Unlike separate rawinsondes, the SIRS instrument tends to make all errors in the same direction, thus not significanly affecting gradients.

C. GRADIENTS OF N_i

This approach is somewhat more direct than the "D" value approach. Here, the derivatives $\frac{1}{f}\frac{\partial N_i}{\partial n}$, were compared directly with the calculated values of V_B and the resulting regression equations were used to calculate V_B for the independent sample. The regression equation derived from the December data is given below as an example.

$$\begin{aligned} V_{BN} = & -2.371x_1 - 11.456x_2 + 30.703x_3 - 1.548x_4 \\ & +64.865x_5 - 25.090x_6 - 2.122x_7 + 2.251x_8 + 24.42 \end{aligned} \tag{19}$$

Here, x_i indicates $\frac{1}{f}\frac{\partial N_i}{\partial n}$.

This approach was tried with data interpolated to evenly spaced grid points and also with the derivatives calculated at the mid-points between the actual, uninterpolated satellite subpoints. In view of the results obtained during the first phase of the experiment, no

data filtering was attempted. The results of this approach are summarized in Table IV. Tests of the December 30 and January 5 independent samples for the in situ SIRS data were not made in view of the more favorable results for the dependent data when subtrack interpolation was employed.

The results here are somewhat mixed. Only for the December 30 case does the scheme show as much accuracy as the results for the "D" value approach shown in Table III. The differences between the results of this method and those of the "D" value approach may be explained by the greater number of derivatives which were taken in the $\frac{1}{f}\frac{\partial N_i}{\partial n}$ approach. In the latter case, a derivative was calculated for each SIRS channel, while in the "D" value approach the only derivative calculated was that of the ballistic "D" value. The smaller explained variance with the $\frac{1}{f}\frac{\partial N_i}{\partial n}$ approach probably resulted from a summation of the roundoff errors in each derivative calculation. In this approach, interpolation to obtain uniform spacing of data produced a greater improvement in results than it did in the "D" value approach. This is perhaps explainable by the very small values of $\frac{\partial N_i}{\partial n}$ compared to the magnitudes of $\frac{\partial D_B}{\partial n}$. With such small values, the loss of definition due to the wide spacing of some of the uninterpolated data could seriously degrade results. The improvement resulting from interpolation is indicative of the cubic spline fitting routine's ability to accurately determine the data patterns.

TABLE IV

Comparison of Rawinsonde-Derived and Regression Equation-Derived Winds Using Gradient of N_i Approach (SIRS-A)

Sample	n	Mean V_B	Std Dev of V_{BN}	Mean V_{BN}	Corr Coeff	Explained Variance	Std Error	Max Error
Interpolated								
27-29 Dec	68	5.6kts	18.9kts	16kts	.809	65.4%	11.9kts	24.4kts
30 Dec	27	6.6	28.5	24	.956	91.4	8.5	24.1
2-4 Jan	74	35	13.2	35	.677	45.9	13.0	39.5
5 Jan	19	37	13.1	37	.626	39.1	10.5	24.1
Uninterpolated								
27-29 Dec	45	7.3	18.4	16.6	.530	28.0	17.3	33.3
2-4 Jan	60	29	17.4	29.0	.347	12.0	17.3	43.0

D. DIRECT APPROACH

A third method tried was the derivation of a regression equation by direct correlation between V_B and N_i. This approach gave poorer results than the previous two methods used, as Table V indicates. (See also Tables III and IV.)

TABLE V

Comparison of Winds Obtained From Direct Approach with Rawinsonde-Derived Winds (SIRS-A)

Sample	n	Corr Coeff	Expalined Variance	Standard Error	Max Error
27-29 Dec	80	.788	62.2%	12.8kts	34.3kts
30 Dec	31	.771	59.5	18.0	31.1
2-4 Jan	88	.607	36.8	15.6	38.8
5 Jan	23	.372	13.8	13.3	21.1

Except in the case of the December 27-29 "D" value approach, these results are somewhat inferior to those obtained by either of the first two methods described. Two factors could account for the superior performance in December. First, the value of $\overline{W}_B$ is smaller in December which would tend to make smaller errors. Secondly, the patterns obtained on the u and v analyses were more constant with time in the December case and thus the regression coefficients obtained from the dependent sample would be more applicable to the independent data.

E. SUMMARY

Both the "D" value approach and the gradient of N_i approach appear capable of providing usable values of V_B. The direct approach, however, in most cases gave correlations too low to be considered of value. It is interesting to note that the correlations for the independent data sets were higher than for the dependent data with the December data. This result was also noted by Elsberry and Martin [2] in their use of the same data for ballistic density calculations.

To further check the temporal stability of the techniques, the December equations were applied to the January data set and vice versa. The resulting values of ballistic wind were then compared to the rawinsonde-derived values. Table VI gives a summary of the results for the "D" value approach. All runs were performed using using equally spaced data.

TABLE VI

Results of Temporal Stability Checks of the "D" Value Approach

Sample	Corr Coeff	Explained Variance	Std Error
Dec Data Jan Equations			
$V1-V_{BN}$	.123	1.5%	21.0kts
$V_1-V_{BN}(I)$,Dec 30	.677	45.9	22.0
Jan Data Dec Equations			
$V1-V_{BN}$	.413	17.1	13.1
$V1-V_{BN}(I)$,Jan 5	.370	13.7	11.2

It can be seen that the temporal stability is rather poor. The percentages of explained variances were generally so low as to be insignificant. It is interesting to note that the best performance was for the December 30 data using equations formed from the later data set. A possible implication to be drawn from Table VI is that the regression equations would have to be continually updated using data from a period immediately prior to the day for which the independent wind values are to be calculated.

In the tables of results for the various methods, it may be noticed that the sample sizes vary considerably although the samples are taken from the same data set. This is the result of the loss of end points from the rows of data in the finite differencing processes used to calculate derivatives.

IV. SIRS-B EXPERIMENT

A. DATA AND DATA HANDLING

The SIRS-B data used in the second phase of experimentation were taken from the period January 10-17, 1971. Data were selected from the latitude belt from 40N to 75N. As was the case in previous experimentation, corrected clear column radiances were used. Concurrent rawinsonde data were also obtained. Though it was expected to sacrifice some data density to obtain the side-looking capability of the SIRS-B, data coverage was still generally poor. Data points were widely scattered and many areas of good rawinsonde coverage were completely devoid of the SIRS-B data. In areas with SIRS-B coverage, the normal spacing between data points was 450 to 650 kilometers. Another shortcoming of the SIRS-B data was that the window channel was inoperative so that only seven channels were available.

To permit the calculation of gradients, the data were chosen from areas where several SIRS data points could be found grouped in a "data cluster." These clusters consisted of one master data point surrounded by three to six other "slave" points. A schematic of a data cluster showing a "master" and five "slaves" is shown in Fig. 2. The maximum distance between a master and a slave was arbitrarily limited to 1000 kilometers. The surface between each master-slave pair was assumed to be linear and a plane was fitted between each pair. The average slope for all of the master-slave

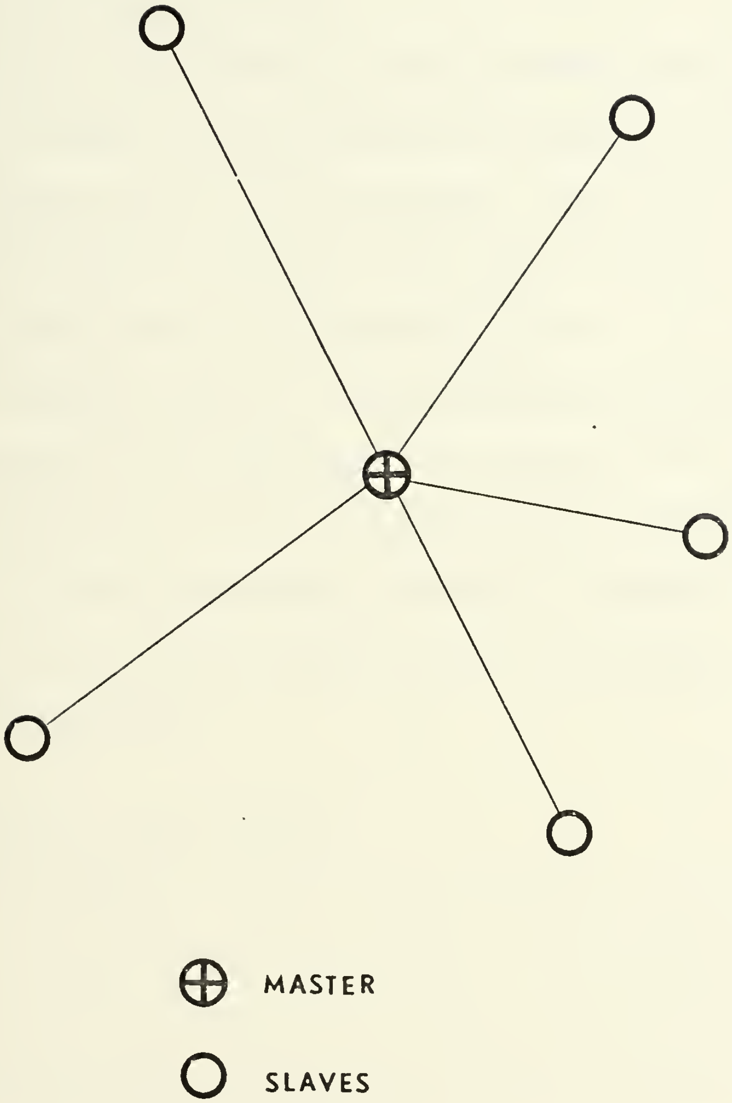

Fig. 2 Sample SIRS-B Data Cluster

pairs was then taken as the partial derivative with respect to the x or y direction for either the "D" value or the radiance value at the master point.

For the period January 10-17, 45 separate data clusters were found in areas with good rawinsonde coverage. These clusters comprised a total of 194 data points. The first 30 clusters covering January 10-14 were used as the dependent sample in all experiments and the 15 clusters from January 15-17 were used as the independent sample. The side-looking capability of the SIRS-B instrument enabled calculation of both the u and v components of the ballistic wind. Eqs. (11) and (13) were therefore broken down into their apprcpriate components. Otherwise, computational procedures used were essentially the same as those used in the SIRS-A experiment.

B. THE "D" VALUE APPROACH

The regression equation for calculation of ballistic "D" values was obtained by correlation of rawinsonde-derived "D" values with the radiance data for all 130 data points in the dependent sample. The resulting equation is given as Eq. (20).

$$D_B = -10.82N_2 + 15.23N_4 - 47.41N_6 + 95.96N_7 - 30.45N_8 - 1727.6 \qquad (20)$$

Note that in this equation, channels 3 and 5 were not statistically significant.

The correlation coefficient between "D" values and radiances for the dependent sample was .906, indicating 81.2 per cent of the variance was explained. The standard error was 119.8 meters. Though this standard error was quite large, no data filtering was attempted due to the already small sample size and the negligible improvements with filtering in the SIRS-A experiment.

"D" values obtained from the application of Eq. (20) to the data points in the indpendent sample were next used in Eq. (11) to obtain u and v components of ballistic wind. These SIRS-B derived ballistic winds were then correlated with rawinsonde ballistic winds to check the accuracy of this method. The results of these correlations for the independent sample are tabulated in Table VII.

It can be seen that the "D" value approach does quite well at determining the v component of $\mathbb{V}_B$. In fact, the results are comparable to those for the one-day January dependent sample as noted in Table III. The poorer results with the u component may have resulted from the relatively small population of sample clusters. The method used to calculate gradients, when used over the wide spacings involved in a typical cluster might also contribute to the low correlations.

C. GRADIENTS OF N_i

This technique, using Eq. (13), was carried out in essentially the same manner as the experiment with SIRS-A data. In this case, however, the u and v components were correlated with values of $\frac{1}{f}\frac{\partial N_i}{\partial y}$ and $\frac{1}{f}\frac{\partial N_i}{\partial x}$ respectively.

The resulting equations for these components, u_B and v_B are given below.

TABLE VII

Comparison of Ballistic Winds Derived
Using "D" Value Approach with Observed
Ballistic Winds (SIRS-B)

	n	Mean V_B	Std Dev of V_B	Mean $\|V_B\|$	Corr Coeff	Expl'd Variance	Std Error	Max Error
u component	15	21kts	19kts	24kts	.525	27.6%	17kts	32kts
v component	15	-2.5	16	14	.807	65.2	9.5	18

$$u_B = 152.2x_2 - 255.8x_3 + 344.7x_4 + 197.6x_6 - 695.7x_7 + 292.7x_8 + 14.74 \tag{21}$$

$$v_B = -133.4y_2 - 121.0y_3 + 144.4y_4 - 146.0y_6 + 568.8y_7 - 129.6y_8 + 5595. \tag{22}$$

As before, the x_i represent $\frac{1}{f}\frac{\partial N_i}{\partial y}$ and the y_i represent $\frac{1}{f}\frac{\partial N_i}{\partial x}$. The results using these equations to calculate u_B and v_B are given in Table VIII. Again, the regression technique gave relatively poor results for the calculation of u_B in the independent sample. The high correlations for u_B in the dependent sample indicate that the low correlation for the independent sample may be due to data shortcomings as previously indicated. Comparison with Table IV shows that the results for v_B are better than for the January SIRS-A case. As was noted in the SIRS-A experiment, this method explained a somewhat lower percentage of the variance than did the "D" value approach.

D. DIRECT APPROACH

Direct correlations of the seven SIRS channel radiances with both u_B and v_B were calculated and the resulting regression equations were used as with the SIRS-A data. The results are tabulated in Table IX. Note that in this approach the sample sizes are much larger than in either of the other methods.

This method also specifies u_B much less accurately than it does v_B. The results for both components are not as good as those obtained with the SIRS-A data. (See Table V.)

TABLE VIII

Comparison of Rawinsonde-Derived and Regression Equation-Derived Winds Using N_i Approach (SIRS-B)

Dependent Sample	n	Mean V_{BN}	Std Dev of u_B, v_B	Mean $\lvert V_{BN} \rvert$	Corr Coeff	Expl'd Variance	Std Error	Max Error
u component	30	23kts	23kts	27kts	.735	54.0%	18kts	-29kts
v component	30	-4.6	25	20	.764	58.4	18	-37
Independent Sample								
u component	15	21	19	24	.529	28.0	17	32
v component	15	-2.4	15	14	.722	52.1	11	29

TABLE IX

Comparison of Ballistic Winds Derived by Direct Approach with Rawinsonde-Derived Winds (SIRS-B)

Dependent Sample	n	Corr Coeff	Explained Variance	Std Error	Max Error
u component	130	.288	8.3%	21kts	-47kts
v components	130	.535	28.6	20	-52
Independent Sample					
u component	64	.302	9.1	17	41
v component`	64	.452	20.4	13	30

Because of the wide geographical distribution of the SIRS-B data the data sample is much more heterogeneous. The results of this test seem decisively insignificant.

E. SUMMARY

In neither the u_B nor v_B case were the results of the SIRS-B experiments as good as those obtained with the December SIRS-A data. For the v component the results were considered quite good considering the lack of window channel data and the wide spatial and temporal distribution of the data. Since all three approaches gave significantly poorer results for the u components, it is thought that a major portion of the difficulty must lie either in the nature of the u field or in the small size of the independent sample.

When correlations between the rawinsonde-derived ballistic winds and winds geostrophically calculated from analyzed rawinsonde-derived

"D" values were computed, the results were mixed. For the entire test period ($n = 45$) the correlation coefficient for u_B was .848 and for v_B was .815. Making a similar test of correlation using only the data of Jan. 15-17 ($n = 15$), however, the correlation coefficient for u_B was .812 and for v_B was .431. These anomalous results indicate the variability of the data in the period of the SIRS-B experiment and the need for more closely spaced SIRS data points in future experimentation.

V. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

Both the "D" value and gradient of radiance approaches were found to be capable of specifying the ballistic wind within the error limits set down by Finke [3] . Generally, the "D" value approach explained a higher percentage of the variance and gave lower standard errors.

Contributing factors to the generally poor results for the u components upon independent data testing in the SIRS-B experiment may have been one or a combination of the following. The nature of the u_B field could be such that it is not well specified by a linear regression equation. It was noted that the regression equations usually failed to calculate any negative values of u_B. In only one of the two cases where negative values were calculated was the u_B obtained from rawinsondes also negative. Also, the previously exhibited poor temporal stability of the regression equations could be more pronounced for u_B in this data sample.

The direct approach is not thought to be capable of claculating ballistic winds accurately enough to be useful.

B. RECOMMENDATIONS

It is recommended that experimentation be continued with a new SIRS-B data set having a more even data distribution around the northern hemisphere. Better data coverage would enable the use of shorter periods of time to derive the equations which would help

counteract the temporal stability problem. Additionally, rocketsonde data should be used in addition to rawinsonde data to more accurately define the motion and mass fields at upper levels. This, however, tends to limit data to the northern hemisphere.

Experimentation with a long data record applied to homogeneous latitude should be conducted to determine the optimum time period for determining the regression equations. An optimum record length should be found which balances sample size and representativeness of the early portions of the record.

In further SIRS-B experimentation, a more sophisticated method for calculation of gradients should be utilized. One method suggested by Kung [4] is to find the least squares approximation to the solution of a system of equations of the following form:

$$\Delta Q_i = \frac{\partial Q}{\partial x}\Delta x_i + \frac{\partial Q}{\partial y}\Delta y_i \qquad (23)$$

Here Q is any scalar variable, and Δx and Δy are the eastward and northward distances between the i^{th} pairing of the master data point and a slave. The unknown partial derivatives may be calculated if $i \geq 2$ in each master-slave cluster. The greater the degree of overspecification, the better should be the least squares fit.

Another method for obtaining gradients might be the description of the scalar surface as a polynomial by the   of a regression equation in x and y. Derivatives could readily be calculated from the polynomial. This method would be time consuming since a separate regression equation would have to be generated for each set of points.

LIST OF REFERENCES

1. Dixon, W. J., Biomedical Computer Programs, Los Angeles, Health Sciences Computing Facility, University of California, Los Angeles, p. 585, 1966.

2. Elsberry, R. L., and F. L. Martin, "An Experimental Method of Determining Ballistic Densities Making Direct Use of SIRS Radiances," Rept. No. NPS-51ES,MR71101A, Naval Postgraduate School, p. 18, 1971.

3. Finke, R. G., "Reentry Vehicle Dispersion Due to Atmospheric Variations," Research Paper P-506, Institute for Defense Analysis Science and Technology Division, p. 42, 1969.

4. Kung, E. C., "Kinetic Energy Dissipation in the Large-Scale Atmospheric Circulation," Mon. Wea. Rev., v. 94, p. 67-82, 1966.

5. Ruggles, K. W., Private Communication, 1970.

6. Smith, W. L., Woolf, H. M., and Jacob, W. J., "A Regression Method for Obtaining Real-Time Temperature and Geopotential Height Profiles from Satellite Spectrophotometer Measurements and its Application to Nimbus III "SIRS" Observations," Mon. Wea. Rev., v. 98, p. 582-603, 1970.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Dr. R. L. Elsberry Department of Meteorology, Code 51Es Naval Postgraduate School Monterey, California 93940	6
4.	LT John Wright La Belle Missouri 63447	2
5.	Dr. F. L. Martin Department of Meteorology, Code 51Mr Naval Postgraduate School Monterey, California	1
6.	Officer-in-Charge Naval Weapons Engineering Support Activity Detachment (FAMOS) 3737 Branch Avenue, Room 307 Hillcrest Heights, Maryland 20031	1
7.	Mr. Paul Lowe Environmental Prediction Research Facility Naval Postgraduate School Monterey, California 93940	1
8.	Department of Meteorology, Code 51 Naval Postgraduate School Monterey, California 93940	1

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

ORIGINATING ACTIVITY (Corporate author)

Naval Postgraduate School
Monterey, California 93940

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

REPORT TITLE

Regression Techniques for the Calculation of Ballistic Winds Utilizing Satellite Infrared Spectrophotometer Data

DESCRIPTIVE NOTES (Type of report and, inclusive dates)

MASTER'S THESIS; September 1971

AUTHOR(S) (First name, middle initial, last name)

John R. Wright

REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1971	43	6

. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
. PROJECT NO.	
.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
.	

. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Postgraduate School Monterey, California 93940

. ABSTRACT

Three regression techniques for the calculation of ballistic winds using satellite infrared spectrophotometer (SIRS) data are evaluated. Experiments testing the derived methods are carried out using two SIRS-A data sets and one SIRS-B data set. In all three cases, the data sets are divided into a large dependent sample used to derive the regression equations and smaller independent sample for testing the equations' validity. One method, that of geostrophic calculations using regression derived "D" values of ballistic geopotential height is found to be superior to the other methods.

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Ballistic Winds						
Satellite Infrared Spectrophotometer Radiances						